

ESG-Ratings: Signale ohne Substanz?

Warum Aufsichtsräte ESG-Ratings kritisch prüfen sollten

Unternehmen nutzen ESG-Ratings zur Steuerung der externen Wahrnehmung ihrer Nachhaltigkeitsperformance. Aufsichtsräte sollten ESG-Ratings daher nicht unkritisch verwenden. Eine aktuelle Studie der Universität Göttingen analysiert für europäische Unternehmen die Verbreitung der verschiedenen ESG-Ratings und zeigt Fallstricke bei ihrer Interpretation auf.

Max Wendler, Michael Wolff

I. Einleitung

Verschiedenste externe Stakeholder wie Investoren oder NGOs verlangen zunehmend Informationen über die Nachhaltigkeitsaktivitäten und -performance von Unternehmen. Zur Messung und vor allem zur externen Kommunikation von Nachhaltigkeitsaspekten nutzen Unternehmen regelmäßig ESG-Ratings, z.B. im Rahmen von Investorengesprächen. Vor diesem Hintergrund sollten Aufsichtsräte ein Verständnis dafür entwickeln, wie ESG-Ratings entstehen und was diese messen, um mögliche Implikationen für die Unternehmensführung richtig einordnen zu können.

In der Praxis existieren zahlreiche Ansätze bzw. Methodiken zur Ableitung von ESG-Ratings, mit teils deutlich unterschiedlichen Zielsetzungen. Gleichzeitig wird zunehmend Kritik an den Ratings geäußert, da Bewertungen häufig inhaltlich nicht nachvollzogen werden können und auch der Zusammenhang zwischen der Ratingausprägung und der tatsächlichen Nachhaltigkeitsperformance unklar bleibt. In der Folge ist das Vertrauen in ESG-Ratings unter Unternehmensvertreterinnen und -vertretern gering, wie die Umfrageergebnisse des letzten „Rate the Raters“-Reports des SustainAbility Institute zeigen. Diese zeigen jedoch auch, dass Stakeholder, z.B. Investoren, Ratings zunehmend höhere Bedeutung beimessen – was auch am Mangel an besseren Alternativen liegen mag. Für Unternehmen eröffnet dies die Möglichkeit, die Undurchsichtigkeit der bestehenden Rating-Landschaft zu nutzen und ausgewählte Ratings gezielt als Signale einzusetzen, um sich in ein möglichst positives Licht zu rücken. Um einen informierten Diskurs mit dem Management führen und ihre Kontrollfunktion effektiv erfüllen zu können, sollten Aufsichtsratsmitglieder die bestehenden Unzulänglichkeiten von ESG-Ratings kennen und wissen, dass Unternehmen diese strategisch einsetzen können.

II. ESG-Ratings: Der Status quo

Vor allem drei Kritikpunkte werden an den aktuellen ESG-Ratings geäußert. Erstens wird die Divergenz zwischen Ratings verschiedener Agenturen genannt: Da unterschiedliche – oft proprietäre und daher intransparente – Methodiken eingesetzt werden, kann es vorkommen, dass ein Unternehmen deutlich unterschiedliche Bewertungen von verschiedenen Anbietern erhält (z.B. einen hohen Score von MSCI, aber einen niedrigen von Sustainalytics). Laut einer umfassenden Analyse von Berg et al. (2022) liegt dies neben der Verwendung unterschiedlicher Messmethoden insbesondere an unterschiedlichen Betrachtungsbereichen – also der Frage, welche Aspekte der ESG-Performance in einem Rating berücksichtigt werden. Zudem zeigen unterschiedliche Studien, dass durch die mangelnde Standardisierung des unternehmerischen ESG-Reportings – die wichtigste Datenquelle vieler Ratings – und die manuelle Bewertung durch Analysten dieses Problem noch verstärkt wird, da Subjektivität und systematische Verzerrungen in die Bewertungen Einzug halten.

Als zweiter Kritikpunkt gilt die Vermischung verschiedener Perspektiven. Grundsätzlich gibt es zwei unterschiedliche Ansätze zur ESG-Bewertung: ESG-Risiko-Ratings bewerten die Exposition von Unternehmen gegenüber wirtschaftlichen ESG-Risiken (z.B. MSCI, Sustainalytics), während ESG-Impact-Ratings die Auswirkungen von Unternehmen auf Umwelt und Gesellschaft messen (z.B. Thomson Reuters Refinitiv). Um falsche Schlussfolgerungen zu vermeiden, ist es wichtig, die jeweiligen Ziele und Methoden hinter verschiedenen ESG-Ratings zu verstehen.

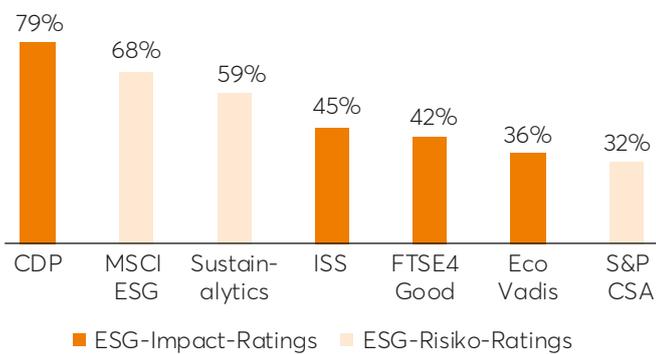
Drittens darf bezweifelt werden, dass ESG-Ratings tatsächlich aussagekräftig sind hinsichtlich dessen, was sie zu messen vorgeben: Aktuelle Analysen zeigen, dass ESG-Ratings zwar mit dem Umfang des Reportings eines Unternehmens

korrelieren, aber nicht mit „harten“ Performance-Indikatoren wie CO₂-Emissionen oder Gesetzesverstößen.

III. ESG-Ratings: Verwendung durch Unternehmen

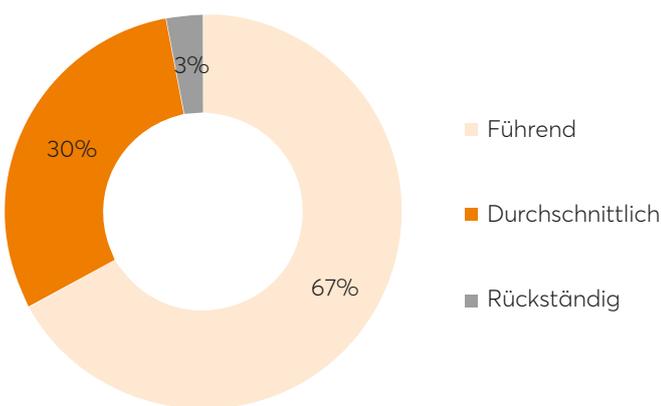
Eine aktuelle Studie der Universität Göttingen untersucht, welche ESG-Ratings große börsennotierte europäische Unternehmen (Mitglieder des STOXX Europe 600 Index) öffentlich kommunizieren – vor allem auf Webseiten und in Nachhaltigkeitsberichten. Dabei zeigt sich, dass die Mehrheit (88%) solcher Unternehmen sich mittlerweile öffentlichkeitswirksam mit ESG-Ratings schmückt: Im Median kommunizieren diese Unternehmen vier verschiedene Ratings. Am häufigsten werden Ratings etablierter, bekannter Agenturen verwendet.

Abb. 1: Häufigste kommunizierte ESG-Ratings



Wie Abb. 1 zeigt, werden die drei meistverwendeten Ratings von jeweils mehr als der Hälfte der Unternehmen genannt, Ratings des Carbon Disclosure Project (CDP) sogar von fast 80% der Unternehmen. Dies passt zu den Umfrageergebnissen des „Rate the Raters“-Reports, nach denen sowohl Investoren als auch Unternehmen dieses Rating als das mit Abstand qualitativ hochwertigste und aussagekräftigste ansehen.

Abb. 2: Kommunizierte Ratings nach Performance



Da Unternehmen frei entscheiden können, ob sie Bewertungen, die sie erhalten, selbst öffentlich kommunizieren, liegt nahe, dass sie nur solche Ratings veröffentlichen, in denen sie gut abschneiden – insbesondere angesichts der teils erheblichen Divergenz zwischen verschiedenen Agenturen. Tatsächlich kommt die Studie der Universität Göttingen zu dem Ergebnis, dass 97% aller Ratings, für die Unternehmen einen konkreten Score angeben, einer mindestens durchschnittlichen Performance entsprechen, mehr als zwei Drittel sogar einer führenden (s. Abb. 2).

Dieses Muster setzt sich fort, wenn man zusätzliche Informationen betrachtet, die Unternehmen als Referenz angeben, insbesondere hinsichtlich der Entwicklung ihrer Leistung im Vergleich zu einer früheren Bewertung oder hinsichtlich des Vergleichs ihrer Leistung mit dem jeweiligen Industriedurchschnitt: In 96% der Fälle, in denen sich Unternehmen in ihrer Industrie vergleichen, geben sie an, dass ihre eigene Performance besser ist als der Durchschnitt ihrer Referenzgruppe. Und in 83% der Fälle, in denen Unternehmen ihr aktuelles Rating mit einem früheren Rating vergleichen, ist die neue Bewertung mindestens so hoch wie die alte, in 39% ist sie sogar besser. Zum Vergleich: Bei einer Analyse der jährlichen Veränderung des Refinitiv-ESG-Score von Thomson Reuters für die Unternehmen in der Stichprobe verbesserten oder verschlechterten sich jeweils 40–50% der Unternehmen – negative Veränderungen werden also mehrheitlich nicht angegeben.

Aus Abb. 1 ist zudem ersichtlich, dass unter den meistverwendeten Ratings zu etwa gleichen Teilen Impact- und Risiko-Ratings zu finden sind. Dies ergibt im Hinblick auf die eingangs diskutierten unterschiedlichen Interessen verschiedener Stakeholder zwar Sinn, jedoch wird in mehr als 60% der Fälle nicht erklärt, aus welcher Perspektive ein ESG-Rating ermittelt wird. Dies kann zu Verwirrung führen. Für eine effektive Diskussion der ESG-Performance eines Unternehmens ist es maßgeblich, dass allen Beteiligten klar ist, ob die diskutierte Bewertung finanzielle Risiken oder Auswirkungen auf Umwelt und Gesellschaft repräsentiert.

Weiterführende Analysen liefern zudem Hinweise darauf, dass ESG-Ratings selbst sowie der Umfang, in dem ein Unternehmen diese kommuniziert, nur limitierte Aussagekraft hinsichtlich der tatsächlichen ESG-Performance eines Unternehmens haben. Insgesamt legen die Ergebnisse nahe, dass sowohl die Güte erhaltener ESG-Scores als auch die Anzahl der Ratings, die ein Unternehmen kommuniziert, vor allem damit korrelieren, wie viel Ressourcen dieses in den Aufbau seines ESG-Reportings investiert – nicht jedoch damit, wie viel Geld, Zeit und Energie in konkrete Initiativen zur Verbesserung der Auswirkungen auf Umwelt und Gesellschaft (z.B. in Maßnahmen zur Emissionsreduktion) fließen.

Zum einen bestätigen auch die Ergebnisse der aktuellen Göttinger Studie die in vorherigen Studien aufgezeigte systematische Verzerrung zugunsten großer Unternehmen. Die nach Marktkapitalisierung größten Unternehmen in der aktuellen Stichprobe haben einen um zehn Punkte höheren Refinitiv-ESG-Score und kommunizieren im Schnitt fast ein Rating mehr als durchschnittlich große Unternehmen. Eine naheliegende Erklärung ist, dass höher bewertete – und damit stärker im Rampenlicht stehende – Unternehmen unter höherem Rechtfertigungsdruck stehen, aber auch mehr verfügbare Ressourcen haben, um diesem durch Investment in ihr ESG-Reporting zu begegnen.

Zum anderen veranschaulichen die in Tab. 1 dargestellten Ergebnisse einen ähnlichen Zusammenhang für Unternehmen mit hoher Emissionsintensität. Während man intuitiv erwarten würde, dass dies in Ratings abgestraft wird, haben

Tab. 1: Anzahl Ratings nach Emissionsintensität

Unternehmensperzentile nach Emissionsintensität	Anzahl Firmen	Ø Anzahl kommunizierter Ratings	Ø ESG-Score (TR Refinitiv ESG)
Höchste (>90%)	48	5,0	73,2
Höher (65-90%)	120	4,6	73,1
Durchschnittlich (35-65%)	144	4,2	72,4
Geringer (10-35%)	120	4,0	69,6
Geringste (<10%)	49	3,1	62,4

die Unternehmen mit der höchsten Emissionsintensität sogar einen im Schnitt mehr als zehn Punkte besseren Score als jene mit der niedrigsten Emissionsintensität – und kommunizieren fast zwei Ratings mehr (und es werden nur positive Scores veröffentlicht). Auch hier liegt nahe, dass der spezifische Rechtfertigungsdruck eines hohen Emissionsniveaus mit einem erhöhten Investment in das ESG-Reporting einhergeht.

Tab. 2: Grünes Humankapital und ESG-Ratings

Einflussfaktoren (Auszug aus Korrelationsanalyse)	Anzahl kommunizierter Ratings
Investment in „Compliance-bezogenes“ grünes Humankapital	¼
Investment in „substantielles“ grünes Humankapital	¾

Drittens zeigen die in Tab. 2 veranschaulichten Ergebnisse einer Korrelationsanalyse der Einflussfaktoren zwischen dem Investment in spezifisches Humankapital (auf Basis der in Stellenanzeigen geforderten Qualifikationen) und der Anzahl der von Unternehmen öffentlich kommunizierten Ratings einen Zusammenhang für „Compliance-bezogenes“, jedoch nicht für „substantielles“ grünes Humankapital. Dabei umfasst „Compliance-bezogenes“ grünes Humankapital Qualifikationen und Wissen rund um Umweltvorschriften und Berichtsstandards, während „substantielles“ grünes Humankapital die notwendigen Management- und technischen Fähigkeiten zur Planung, Durchführung und Steuerung von Umweltinitiativen kombiniert, die auf konkrete operative Prozessverbesserungen gerichtet sind.

IV. Implikationen für Aufsichtsräte

Aufsichtsratsmitglieder sollten sich über die Limitationen von ESG-Ratings und den sich daraus ergebenden Möglichkeiten zur strategischen Kommunikation durch das Management der von ihnen überwachten Unternehmen im Klaren sein. Sie sollten sich nicht nur auf Ratings verlassen, die ihnen vom Management vorgelegt werden, sondern aktiv mehrere Ratings vergleichen – und dabei hinterfragen, welche Indikatoren in die Bewertung einfließen und ob die jeweilige Methodik für eine Bewertung der spezifischen Unternehmenssituation passend und geeignet ist. Insbesondere die Divergenzen zwischen verschiedenen Anbietern und die Tatsache, dass Ratings oft mehr mit Reporting-Aktivitäten als mit substantiellen Verbesserungen korrelieren, erfordern eine offene Diskussion über die Grenzen dieser Metriken.

Um eine effektive Diskussion über ESG-Ziele und deren Erreichung gemeinsam mit dem Management zu führen,

gilt es, zunächst klare Vorstellungen über die ESG-Strategie zu entwickeln: Welche ESG-Themen sind von materieller Bedeutung für das Unternehmen – aus einer Impact- und/oder einer Risiko-Perspektive? Ratings können in dieser Diskussion eine Rolle spielen, jedoch sollten diese weniger als Ziel oder Selbstzweck, sondern eher als Stichwortgeber für mögliche Diskussionsthemen gesehen werden. Wenn ein gemeinsames Verständnis hergestellt ist, können im nächsten Schritt konkrete, messbare Ziele formuliert und die richtigen Metriken zu ihrer Nachhaltigkeit ausgewählt werden. Dies können bestimmte Ratings sein – da diese aber oft viele verschiedene Indikatoren einbeziehen, von denen nicht alle für das Unternehmen gleichermaßen relevant sein müssen, sollte darüber nachgedacht werden, sich stattdessen auf einige wenige spezifische Kennzahlen zu konzentrieren, die auf die gemeinsam formulierten Ziele abgestimmt sind.

V. Fazit

Die Aussagekraft von ESG-Ratings ist begrenzt. Methodische Divergenzen und unterschiedliche Perspektiven begünstigen eine selektive Kommunikation. Der Bezug zu konkreten Verbesserungen ist gering. Aufsichtsratsmitglieder sollten Ratings, die das Management vorlegt, kritisch prüfen und hinterfragen, welche die maßgeblichen Kennzahlen sind, die die tatsächliche Performance des Unternehmens hinsichtlich materieller ESG-Themen widerspiegeln.

»AR1478344

Literaturhinweise:

- Berg/Köbel/Rigobon, Aggregate Confusion: The Divergence of ESG Ratings, *Review of Finance*, 26(6) 2022 S. 1315-1344.
- The Sustainability Institute by ERM, Rate the Raters 2023: ESG Ratings at a Crossroads, März 2023.

Zusammenfassung der Studienerkenntnisse

1. Unternehmen kommunizieren nahezu ausschließlich positive ESG-Ratings und Referenzpunkte.
2. Unternehmen erklären für Ratings, die sie kommunizieren, mehrheitlich nicht, aus welcher Perspektive diese bewerten (finanzielles Risiko vs. Auswirkungen auf Umwelt und Gesellschaft).
3. Weiterführende Analysen werfen Zweifel an der generellen Aussagekraft von ESG-Ratings und deren Verwendung in der Unternehmenskommunikation auf.



Max Wendler ist Doktorand am Lehrstuhl für Management und Controlling der Georg-August-Universität Göttingen.



Prof. Dr. Michael Wolff ist Professor für Management und Controlling an der Georg-August-Universität Göttingen.